

**Long Memory Process in Asset
Returns with Multivariate GARCH
Innovations**

Imène Mootamri

June 2011

DT-GREQAM

Long Memory Process in Asset Returns with Multivariate GARCH innovations

Imène Mootamri*

June, 2011

Abstract

The main purpose of this paper is to consider the multivariate GARCH (MGARCH) framework to model the volatility of a multivariate process exhibiting long term dependence in stock returns. More precisely, the long term dependence is examined in the first conditional moment of US stock returns through multivariate ARFIMA process and the time-varying feature of volatility is explained by MGARCH models. An empirical application to the returns series is carried out to illustrate the usefulness of our approach. The main results confirm the presence of long memory property in the conditional mean of all stock returns.

Keywords: Forecasting, Long memory, Multivariate GARCH, Stock Returns.

JEL classification: C22, C51, C53, G12

*GREQAM, Université de la Méditerranée, France, tel: +33.4.91.14.07.70, fax: +33.4.91.90.02.27, Email: imene.mootamri@etumel.univmed.fr

1 Introduction

Long memory seems to be in recent years a very widespread phenomenon in the modelling of economic and financial time series. It can be defined in terms of the persistence of the autocorrelations which decays at a very slow hyperbolic rate. A large numbers of papers demonstrate the existence of long memory in financial economics. Peters (1991) and Greene and Fielitz (1977) found evidence of long-term positive dependence in stock returns by applying the rescaled range (R/S) statistic¹ proposed by Hurst (1951) and modified by Lo (1991). Similar evidence on the German stock return is given by Lux (1996). However, Jegadeesh (1990) challenged the notion of mean reversion for stock returns. He reported negative first order serial correlation and significant positive serial correlation for longer lags using monthly returns for individual stocks. Kim et al. (1991) also challenged the findings of mean reversion. Their findings for post war data showed persistence in returns. Volos and Siokis (2006) examined the presence of long-range dependence in a sample of 34 stock index returns using the procedure of Geweke and Porter-Hudak (1983) and Robinson (1995b). Their results provided significant and robust evidence of fractional dynamics in most major and small stock markets over the sample periods. Goetzmann (1993) applied R/S tests which provided some evidence that the London stock exchange and New York Stock Exchange stock market prices may exhibit long-term memory. Some authors found significant and robust evidence of positive long-term persistence in the Greek stock market (Barkoulas et al. (2000)), in the Brazil stock market (Cajueiro and Tabak (2005) and Cavalcante et al. (2004)), and in the Finnish stock market return data (Tolvi (2003)). Sadique and Silvapulle (2001) examined the presence of long memory in weekly stock returns of seven countries. They found evidence of long-term dependence in four countries. Moreover, Cajueiro and Tabak (2004) found that the markets of Hong Kong, Singapore and China exhibited long-range dependence, while Mills (1993) and Zhuang et al. (2000) investigated British stock returns and found little evidence of long-range dependence. However, Limam (2003) analyzed stock index returns in 14 markets and concluded that long memory tends to be associated with thin markets and Huang and Yang (1999) applied the modified R/S technique to intraday data and provide evidence of long memory phenomenon in both the New York Stock Exchange and Nasdaq indices.

In fact, all these empirical works are based on univariate models. How-

¹The R/S statistic is the range of partial sums of deviations of a time series from its mean, rescaled by its standard deviation.

ever, for many important questions in empirical literature, multivariate settings are preferable. For example, suppose one is considering a portfolio of many assets. The return of the portfolio can be directly computed if one knows the asset shares and the return of each asset (Brooks et al.(2003)). Granger and Joyeux (1980) proposed the univariate fractionally integrated autoregressive moving average (ARFIMA) model to explain the long memory property that exists in the conditional mean. Therefore, we consider in this study the multivariate ARFIMA model², where the fractional integration parameters determine the long memory properties of the data. Apart from the presence of long term dependence in the conditional mean, we will take into account the volatility's properties of the series³. More precisely, the time-varying feature of volatility is explained by multivariate generalized autoregressive conditionally heteroscedastic models (MGARCH).

Despite the fact that univariate descriptions are useful and important, various financial operations require a multivariate framework, since high volatilities are often observed in the same time periods across different assets. The development of MGARCH models from the original univariate specifications represents a major step forward in the modelling of time series: these models permit time-varying conditional covariances as well as variances. To this end, the main contribution of this paper is then to incorporate the MGARCH framework to model the volatility of a multivariate process exhibiting a long term dependence and slow decay in the stock returns. More precisely, we have examined the long memory property in the first conditional moment of daily stock returns; the robustness of the results is also investigated by considering that its innovations are generated by a MGARCH process. We found that the long memory property existed in the conditional mean of the Nasdaq 100, New York Stock Exchange (NYSE) composite and Russell 3000 stock returns.

The rest of this paper is organized as follows. We briefly review the multivariate models in the next Section. Section 3 outlines the Quasi-maximum likelihood estimation and testing procedures for the models which are applied to US stock returns. Section 4 describes multi-step forecasting with some MGARCH model and Section 5 presents the data used and provides empirical results. The paper ends with a short concluding section.

²For more details, see Sowell (1989b) and Gil-Alana (2003-2007).

³See Ding et al.(1993) and Doukhan et al.(2003).

2 Econometric framework

The following is a brief description of the time series models used in this study. The vector ARFIMA models are discussed in detail in Sowell (1989b) and Luceño (1996). Tsay (2007) provided a more detailed description for estimation of the vector ARFIMA models. He suggested a conditional likelihood Durbin-Levinson algorithm to efficiently evaluate the conditional likelihood function of the vector ARFIMA processes. Hosoya (1996) and Nielsen (2004) proposed a class of maximum likelihood estimators and tests. Lobato (1999) analyzed a two-step estimator of the long memory parameters of a vector process by using a semiparametric version of the multivariate Gaussian likelihood function in the frequency domain.

Consider a vector *ARFIMA* processes:

$$\Phi(L) \Delta(L) (y_t - \mu) = \Theta(L) \varepsilon_t \quad (1)$$

where $y_t = (y_{1t}, \dots, y_{nt})'$, $t = 1, \dots, T$, is an n -dimensional vector of observations, $\mu = (\mu_1, \dots, \mu_n)'$ is the conditional mean vector of the process, $\Phi(L) = I_n - \phi_1 L - \dots - \phi_p L^p$ and $\Theta(L) = I_n + \theta_1 L + \dots + \theta_q L^q$ are $(n \times n)$ matrix polynomials in the lag operator L , and satisfy the usual stationary and invertible conditions respectively, i.e., the roots of $\det(\Phi(L))$ ⁴ and $\det(\Theta(L))$ are outside the unit circle. I_n is the identity matrix of order n and $\Delta(L) = \text{diag}[(1 - L)^{d_1}, \dots, (1 - L)^{d_n}]$. The fractional differencing operator is defined by the binomial expansion,

$$(1 - L)^{d_i} = \sum_{j=0}^{\infty} \frac{\Gamma(-d_i + j)}{\Gamma(-d_i) \Gamma(1 + j)} L^j, \text{ for } i = 1, \dots, n$$

where $\Gamma(\cdot)$ is the Gamma function⁵. Hence, the long-range dependence between observations is eventually determined only by the fractional differencing parameter. These characteristics can be seen in the shapes of the spectral density and the autocorrelation function. Indeed, if $|d_i| < 0.5$ for $i = 1, \dots, n$, the multivariate process is stationary and invertible. If $0 < d_i < 0.5$, the process is characterized by strong positive dependency between observations. This is noted in the frequency domain by the spectral density increasing to an infinite value at the zero frequency. In the time domain, the persistence is indicated by the slow decline of the autocorrelation functions which are not absolutely summable. In this case, y_t is said to have a long memory⁶. If

⁴ $\det(\Phi(L))$ is the determinant of the matrix $\Phi(L)$.

⁵The Gamma function is defined as, $\Gamma(t) = \int_0^{\infty} z^{t-1} e^{-z} dz$.

⁶See Beran (1994) for the overviews of long memory processes.

$-0.5 < d_i < 0$, the process exhibits negative dependency between observations. In the frequency domain, this is indicated by the decline of the spectral density to zero, as the frequency approaches zero. The time domain indicates the antipersistence by the rapid decline of the autocorrelation functions.

So, in this paper, we consider a vector ARFIMA model, which generates the long memory property in the first conditional moment and which allows its innovations to be generated by a multivariate MGARCH process. As an illustration, the proposed model is applied to the daily stock returns of Nasdaq, New York Stock Exchange and Russell indices.

The most commonly employed distribution in the literature is the multivariate normal. Thus, we assume that the stochastic vector process $\varepsilon_t = (\varepsilon_{1t}, \dots, \varepsilon_{nt})'$ is conditionally multivariate normal with zero expected value and covariance matrix H_t :

$$\varepsilon_t = H_t^{1/2} z_t, \quad \varepsilon_t | \Omega_{t-1} \sim N(0, H_t) \quad (2)$$

with $E_{t-1}(\varepsilon_t) = 0$. We denoted by Ω_{t-1} the information set generated by the observed series until time $t - 1$. $H_t = (h_{ijt})$ is an $(n \times n)$ conditional covariance matrix of ε_t : $H_t = E_{t-1}(\varepsilon_t \varepsilon_t')$ and z_t is an independent identically distributed random vector error process such that $E(z_t) = 0$ and $E(z_t z_t') = I_n$.

As noted by Silvennoinen and Teräsvirta (2008)⁷, the specification of an MGARCH should be flexible enough to be able to represent the dynamics of the conditional variances and covariances, and, as the number of parameters in an MGARCH model often increases rapidly with the dimension of the model, the specification should be parsimonious enough to allow for easy estimation of the model. Another feature that needs to be taken into account in the specification is that the conditional covariance matrices should be positive definite.

2.1 Generalizations of the univariate GARCH model: VEC and BEKK Models

The *VEC*(p, q) model of Bollerslev et al. (1988) is the first multivariate GARCH model. It is a generalization of the univariate *GARCH* model, where each element of H_t is a linear function of the lagged squared errors as well as cross-products of errors and lagged values of the elements of H_t . The model is given by:

⁷See also Silvennoinen (2008).

$$h_t = C + \sum_{j=1}^q A_j e_{t-j} + \sum_{j=1}^p B_j h_{t-j} \quad (3)$$

where $h_t = \text{vech}(H_t)$ ⁸, H_t is the $(n \times n)$ conditional covariance matrix, $e_t = \text{vech}(\varepsilon_t \varepsilon_t')$, C is a $\left(\frac{n(n+1)}{2}\right)$ parameter vector and, A_j and B_j are $\left(\frac{n(n+1)}{2} \times \frac{n(n+1)}{2}\right)$ parameter matrices.

The main disadvantage of this model is the number of parameters which is equal to $\frac{n(n+1)}{2} \left[(p+q) \left(\frac{n(n+1)}{2} \right) + 1 \right]$ and which becomes larger and larger as the number of variables increases. Thus, the estimation of the parameters is difficult. Furthermore, the positivity of H_t is not guaranteed. To overcome this problem, Bollerslev et al. (1988) suggest the diagonal VEC ⁹ model in which A_j and B_j are assumed to be diagonal, each element h_{ijt} depending only on its own lag and on the previous value of $\varepsilon_{it} \varepsilon_{jt}$. This restriction reduces the number of parameters to $\frac{n(n+5)}{2}$. But even under this diagonality assumption, large scale systems are still highly parameterized and difficult to work with in practice.

One of the most general forms, proposed in Engle and Kroner (1995), is the $BEKK$ (Baba-Engle-Kraft-Kroner) representation. This formulation developed a general quadratic form for the conditional covariance equation which eliminated the problem of assuring the positive definiteness of the conditional covariance matrix. The $BEKK(p, q, K)$ representation for the $(n \times n)$ conditional covariance matrix H_t takes the form:

$$H_t = C'C + \sum_{j=1}^q \sum_{k=1}^K A'_{kj} \varepsilon_{t-j} \varepsilon'_{t-j} A_{kj} + \sum_{j=1}^p \sum_{k=1}^K B'_{kj} H_{t-j} B_{kj} \quad (4)$$

where the summation limit K determines the generality of the process, C is upper triangular $(n \times n)$ matrix, and A_{kj} and B_{kj} are both $(n \times n)$ parameter matrices. This representation guarantees that H_t is positive definite. Although the form of the above model is quite general especially when K is reasonably large, it suffers from the problems due to overparametrization¹⁰.

⁸ vech is an operator that replaces the columns of the lower triangular part of H_t in a $\left(\frac{n(n+1)}{2} \times 1\right)$ vector column.

⁹See also Bauwens et al. (2006) and Silvennoinen and Teräsvirta (2008).

¹⁰See Engle and Kroner (1995) for more discussion on the identification problem of this model.

The parameters of the *BEKK* model do not represent directly the impact of the different lagged terms on the elements of H_t , like in the *VEC* model.

The number of parameters in the *BEKK* model is equal to $(p + q)Kn^2 + \left(\frac{n(n+1)}{2}\right)$ and is still quite large. Thus, as we already mentioned, a problem with these models is that the number of parameters can increase very rapidly as the dimension of the process increases, what creates difficulties in the estimation of the models due to several matrix inversions. So it is typically assumed that $p = q = K = 1$ in applications of this model. A further simplified version of (4) in which A_{kj} and B_{kj} are diagonal matrices has sometimes appeared in applications. This is a diagonal *BEKK* model¹¹ where the number of parameters is equal to $(p + q)Kn + \left(\frac{n(n+1)}{2}\right)$. This model is also a *DVEC* model but it is less general, although it is guaranteed to be positive definite while the *DVEC* is not.

2.2 Nonlinear combinations of univariate GARCH models: CCC and DCC Models

Bollerslev (1990) proposed a class of MGARCH models in which the conditional correlation matrix is time-invariant and thus the conditional covariances are proportional to the product of the corresponding conditional standard deviations. This model is so-called Constant Conditional Correlation (*CCC*). This restriction greatly reduces the number of unknown parameters and thus simplifies the estimation. So, the conditional covariance H_t , may always be decomposed as:

$$\begin{cases} H_t = D_t R D_t \text{ where } h_{ijt} = \rho_{ij} \sqrt{h_{iit} h_{jtt}}, \forall i \neq j \\ D_t = \text{diag}(\sqrt{h_{11t}}, \sqrt{h_{22t}}, \dots, \sqrt{h_{nnt}}) \end{cases} \quad (5)$$

where h_{iit} can be defined as any univariate GARCH model, and R is a $(n \times n)$ symmetric positive definite matrix of conditional correlations with typical element:

$$\rho_{ijt} = \frac{\text{cov}_{t-1}(\varepsilon_{it}, \varepsilon_{jt})}{\text{var}_{t-1}(\varepsilon_{it})^{1/2} \text{var}_{t-1}(\varepsilon_{jt})^{1/2}}$$

with $\rho_{ii} = 1, \forall i = 1, \dots, n$. D_t denotes the $(n \times n)$ diagonal matrix with typical element $\text{var}_{t-1}(\varepsilon_{it})$. The *CCC-GARCH* model assumes that the conditional correlations are constant $\rho_{ijt} = \rho_{ij}$, so that the temporal variation in H_t is determined solely by the time-varying conditional variances for each of

¹¹Bauwens et al. (2006) propose a scalar BEKK model, where A_{kj} and B_{kj} are equal to a scalar times a matrix of ones to reduce also the number of parameters.

the elements in ε_t . As long as each conditional variances are positive¹², the *CCC* model guarantees that the resulting conditional covariance matrices are positive definite.

Despite the simplicity of this model, the assumption that the conditional correlations are constant may seem too restrictive and unrealistic. Engle (2002) proposed a new class of estimator that both preserves the ease of estimation of Bollerslev's constant correlation model yet allows for non-constant correlations. Dynamic Conditional Correlation (*DCC*)-*GARCH*¹³ preserves the parsimony of univariate GARCH models of individual assets' volatility with a simple GARCH process. Further, the number of parameters estimated using maximum likelihood is considerable improvement over both the VEC and the BEKK models. Tse and Tsui (2002) have also proposed a dynamic correlation MGARCH model, however no attempt has been made to allow for separate estimation of the univariate GARCH processes and the dynamic correlation estimator.

The *DCC* model of Engle (2002) computes the time changing conditional correlation matrix from the standardized residuals series:

$$\begin{cases} H_t = D_t R_t D_t \\ D_t = \text{diag}(\sqrt{h_{11t}}, \sqrt{h_{22t}}, \dots, \sqrt{h_{nnt}}) \\ R_t = Q_t^{*-1} Q_t Q_t^{*-1} \end{cases} \quad (6)$$

where $Q_t^* = (\text{diag} Q_t)^{1/2}$, Q_t is the $(n \times n)$ symetric positive definite matrix given by:

$$Q_t = (1 - \theta_1 - \theta_2) \bar{Q} + \theta_1 u_{t-1} u_{t-1}' + \theta_2 Q_{t-1} \quad (7)$$

θ_1 and θ_2 are non-negative scalar parameters satisfying $\theta_1 + \theta_2 < 1$. \bar{Q} is the $(n \times n)$ unconditional covariance matrix composed from the standardized residuals resulting from the first step estimation, where $u_t = (u_{1t}, u_{2t}, \dots, u_{nt})'$ is the standardized residuals vector¹⁴, and

$$Q_t^{*-1} = \text{diag} \left(\frac{1}{\sqrt{q_{11}}}, \frac{1}{\sqrt{q_{22}}}, \dots, \frac{1}{\sqrt{q_{nn}}} \right)$$

The typical element of R_t will be of the form:

$$\rho_{ijt} = \frac{q_{ijt}}{\sqrt{q_{ii}q_{jj}}}$$

¹²See Nelson and Cao (1992) for discussion of positivity conditions for h_{it} in univariate *GARCH*(p, q) models.

¹³Various generalizations of the *DCC-GARCH* model are proposed in the literature (Pelletier (2006), Billio and Caporin (2006), Cappiello et al.(2006), Silvennoinen and Teräsvirta (2009)...).

¹⁴ $u_{it} = \frac{\varepsilon_{it}}{\sqrt{h_{iit}}}$, for $i = 1, \dots, n$.

A slightly different formulation was suggested by Tse and Tsui (2002):

$$R_t = (1 - \theta_1 - \theta_2) R + \theta_1 \Psi_{t-1} + \theta_2 R_{t-1} \quad (8)$$

θ_1 and θ_2 are non-negative scalar parameters such that $\theta_1 + \theta_2 < 1$. Here $R = \{\rho_{ij}\}$ is a time-invariant $(n \times n)$ symmetric positive definite parameter matrix of conditional correlations with ones on the diagonal and Ψ_{t-1} is a $(n \times n)$ correlation matrix of the past m standardized residuals $u_t = (u_{1t}, u_{2t}, \dots, u_{nt})$. The positive definiteness of R_t is ensured by construction if R and Ψ_{t-1} are positive definite such that $m \geq n$. The typical element of Ψ_{t-1} will be of the form:

$$\Psi_{ij,t-1} = \frac{\sum_{l=1}^m u_{i,t-l} u_{j,t-l}}{\sqrt{(\sum_{l=1}^m u_{i,t-l}^2) (\sum_{l=1}^m u_{j,t-l}^2)}} \quad (9)$$

where $u_{it} = \frac{\varepsilon_{it}}{\sqrt{h_{iit}}}$, for $i = 1, \dots, n$.

The number of parameters in both the *DCC* models is equal to $\frac{(n+1)(n+4)}{2}$ if the conditional variances are specified as GARCH(1,1). However, to check if the assumption of the conditional correlations is empirically relevant, one can test $\theta_1 = \theta_2 = 0$. A drawback of the *DCC* models is that θ_1 and θ_2 are scalars, so that all the conditional correlations obey the same dynamics. This is necessary to ensure that R_t is positive definite through sufficient conditions on the parameters.

3 Estimation of MGARCH model

Bollerslev (1990) introduced the CCC-GARCH specification, where univariate GARCH models are estimated for each asset and then the correlation matrix is estimated using the standard closed form MLE correlation estimator by transforming the residuals using their estimated conditional standard deviations. The assumption of constant correlation makes estimating a large model feasible and ensures that the estimator is positive definite, simply requiring each univariate conditional variance to be non-zero and the correlation matrix to be of full rank. However, the constant correlation estimator, as proposed, does not provide a method to construct consistent standard errors using the multi-stage estimation process. Bollerslev (1990) noticed that the notion of constant correlation is plausible but Tse and Tsui (2002) and Tse (2000) found that it can be rejected for some assets.

For the maximum likelihood estimation (MLE) of parameters we assume the conditional normality of the errors. The log-likelihood function of the

model:

$$\begin{cases} \Delta(L)(y_t - \mu) = \varepsilon_t, \varepsilon_t = H_t^{1/2} z_t \\ H_t = D_t R_t D_t \end{cases} \quad (10)$$

has the following form:

$$L_T(\theta) = -\frac{1}{2} \sum_{t=1}^T (n \log(2\pi) + \ln |H_t| + \varepsilon_t' H_t^{-1} \varepsilon_t) \quad (11)$$

where θ is the vector of all the parameters in the model.

Estimating the parameters simultaneously with the conditional variance parameters would increase the efficiency at least in large samples, but this is computationally more difficult. For this reason, we estimate the fractionally integrated model for the conditional mean and we consider $\Delta(L)y_t$ as the data for fitting the MGARCH model, where $\Delta(L) = \text{diag}((1-L)^{\hat{d}_i})$, for $i = 1, \dots, n$.

Engle and Sheppard (2001)¹⁵ showed that the loglikelihood can be written as the sum of a mean and volatility part, depending on a set of unknown parameters and a correlation part that is:

$$\begin{aligned} L_T(\theta) &= -\frac{1}{2} \sum_{t=1}^T (n \log(2\pi) + \ln |H_t| + \varepsilon_t' H_t^{-1} \varepsilon_t) \\ &= -\frac{1}{2} \sum_{t=1}^T (n \log(2\pi) + \ln |D_t R_t D_t| + \varepsilon_t' D_t^{-1} R_t^{-1} D_t^{-1} \varepsilon_t) \\ &= -\frac{1}{2} \sum_{t=1}^T (n \log(2\pi) + 2 \ln |D_t| + \ln |R_t| + u_t' R_t^{-1} u_t) \end{aligned} \quad (12)$$

The conditional variance matrix of a DCC model can be expressed as $H_t = D_t R_t D_t$. The DCC model was designed to allow for two stage estimation, where in the first stage univariate GARCH models are estimated for each residual series, and in the second stage, the residuals, transformed by their standard deviations estimated during the first stage, are used to estimate the parameters of the dynamic correlation. The likelihood used in the first stage involves replacing R_t by the identity matrix in (15). Let $\theta = (\phi_1, \dots, \phi_n, \psi) = (\phi, \psi)$, where the elements of ϕ_i correspond to the parameters of the univariate GARCH model for the i^{th} asset returns, $\phi_i = (c_i, \alpha_i, \beta_i)$, for $i = 1, \dots, n$.

¹⁵See also Sheppard (2003).

The resulting first stage quasi-likelihood function is:

$$\begin{aligned}
Q1L_T(\phi) &= -\frac{1}{2} \sum_{t=1}^T (n \log(2\pi) + 2 \ln |D_t| + \ln |I_n| + \varepsilon_t' D_t^{-1} I_n D_t^{-1} \varepsilon_t) \\
&= -\frac{1}{2} \sum_{t=1}^T (n \log(2\pi) + 2 \ln |D_t| + \varepsilon_t' D_t^{-2} \varepsilon_t) \\
&= -\frac{1}{2} \sum_{t=1}^T \left(n \log(2\pi) + \sum_{i=1}^n \left(\ln(h_{iit}) + \frac{\varepsilon_{iit}^2}{h_{iit}} \right) \right) \\
&= -\frac{1}{2} \sum_{t=1}^T \sum_{i=1}^n \left(\log(2\pi) + \ln(h_{iit}) + \frac{\varepsilon_{iit}^2}{h_{iit}} \right)
\end{aligned} \tag{13}$$

which is simply the sum of the log-likelihoods of the individual GARCH models for each of the asset returns. Once the first stage has been estimated, the second stage is estimated using the correctly specified likelihood, conditioned on the parameters estimated in the first step:

$$\begin{aligned}
Q2L_T(\psi | \hat{\phi}) &= -\frac{1}{2} \sum_{t=1}^T (n \log(2\pi) + 2 \ln |D_t| + \ln |R_t| + \varepsilon_t' D_t^{-1} R_t^{-1} D_t^{-1} \varepsilon_t) \\
&= -\frac{1}{2} \sum_{t=1}^T (n \log(2\pi) + 2 \ln |D_t| + \ln |R_t| + u_t' R_t^{-1} u_t)
\end{aligned} \tag{14}$$

where $u_t = D_t^{-1} \varepsilon_t$ are the univariate GARCH standardized residuals.

Since we are conditioning on $\hat{\phi}$, the only portion of the log-likelihood that will influence the parameter selection is $\ln |R_t| + u_t' R_t^{-1} u_t$, and in estimation of the DCC parameters, it is often easier to exclude the constant terms and simply maximize:

$$Q2L_T^*(\psi | \hat{\phi}) = -\frac{1}{2} \sum_{t=1}^T (\ln |R_t| + u_t' R_t^{-1} u_t) \tag{15}$$

• Test of constant conditional correlations

The first step of modelling time-varying conditional correlations is to test the hypothesis of constant correlations. Testing data for constant correlation has proven to be a difficult problem, as testing for dynamic correlation with data that have time-varying volatilities can result in misleading conclusions and can lead to rejecting constant correlation when it is true due to misspecified volatility models. Bera and Kim (2002) have provided tests of a

null of constant correlation against an alternative of a dynamic correlation structure. It is an information matrix-type test that besides constant correlations examines at the same time various features of the specified model. An alternative test has been proposed by Longin and Solnik (1995). We are interested in testing the null of constant correlation against an alternative of dynamic conditional correlation via the Lagrange Multiplier (LM) approach suggested by Tse (2000) which tested a null of constant conditional correlation against an ARCH in correlation alternative. A rejection of the null hypothesis supports the hypothesis of time-varying correlations. Rewrite the DCC model:

$$\begin{cases} H_t = D_t R_t D_t \\ D_t = \text{diag}(\sqrt{h_{11t}}, \sqrt{h_{22t}}, \dots, \sqrt{h_{nnt}}) \\ R_t = Q_t^{*-1} Q_t Q_t^{*-1} \end{cases}$$

Then, we test

$$H_0 : h_{ijt} = \rho_{ij} \sqrt{h_{iit} h_{jjt}}$$

against the alternative

$$H_A : h_{ijt} = \rho_{ijt} \sqrt{h_{iit} h_{jjt}}$$

where the conditional variances are GARCH (1,1). Under H_0 , the LM statistic is asymptotically $\chi^2(n(n-1)/2)$. Under the normality assumption the conditional log-likelihood of the observation at time t is given by:

$$l_t = -\frac{1}{2} \ln |D_t R_t D_t| - \frac{1}{2} y_t' D_t^{-1} R_t^{-1} D_t^{-1} y_t$$

and the log-likelihood function L is given by $L = \sum_{t=1}^n l_t$.

Engle and Sheppard (2001) proposed another test of the constant correlation hypothesis in the DCC models. The null hypothesis: $R_t = \bar{R} \forall t$ is tested against the alternative: $\text{vech}(R_t) = \text{vech}(\bar{R}) + \lambda_1 \text{vech}(R_{t-1}) + \dots + \lambda_p \text{vech}(R_{t-p})$. The test is easy to implement since H_0 implies that coefficients in the regression $X_t = \lambda_0 + \lambda_1 X_{t-1} + \dots + \lambda_p X_{t-p} + r_t$ are equal to zero, where $X_t = \text{vech}^u(\hat{z}_t \hat{z}_t' - I_n)$, vech^u is like the vech operator but it only selects the elements under the main diagonal, $\hat{z}_t = \hat{R}^{-1/2} \hat{D}_t^{-1} \hat{\varepsilon}_t$ is the $(n \times 1)$ vector of standardized residuals (under H_0), and $D_t = \text{diag}(\sqrt{h_{11t}}, \sqrt{h_{22t}}, \dots, \sqrt{h_{nnt}})$.

• Portmanteau Statistics

It is crucial to check the adequacy of the MGARCH specification. Bollerslev (1990) suggested some diagnostics for the constant correlation MGARCH model. He computed the Ljung-Box portmanteau statistic on the cross products of the standardised residuals across different equations. Critical values were based on the χ^2 distribution. As mentioned by Tse (2000), diagnostics for conditional heteroscedasticity models applied in the literature can be divided into three categories: portmanteau test of the Ljung-Box type, residual-based diagnostic and Lagrange multiplier test. So, to check the overall significance of the residual autocorrelation, we consider the Ljung-Box portmanteau statistic. This test was introduced by Box and Pierce (1970) for goodness-of-fit checking of univariate strong ARMA models. Ljung and Box (1978) proposed a slightly different portmanteau test which is nowadays one of the most popular diagnostic checking tools in ARMA modelling of time series. Following Hosking (1980-1981b), a multivariate version of the Ljung-Box portmanteau¹⁶ statistic is given by:

$$Q_m(k) = T^2 \sum_{j=1}^k (n-j)^{-1} \text{tr} \{ C_{Y_t}^{-1}(0) C_{Y_t}(j) C_{Y_t}^{-1}(0) C_{Y_t}'(j) \}$$

where tr denotes the trace of a matrix. $Y_t = \text{vech}(y_t y_t')$ and $C_{Y_t}(j)$ is the sample autocovariance matrix of order j . Under the null hypothesis, $Q_m(k)$ is distributed asymptotically as $\chi^2(p^2 k)$. Duchesne and Lalancette (2003) generalized this statistic using a spectral approach and obtained higher asymptotic power by using a different kernel than the truncated uniform kernel used in $Q_m(k)$. This test is also used to detect misspecification in the conditional variance matrix H_t , by replacing y_t by $\hat{z}_t = \hat{H}_t^{-1/2} \hat{\varepsilon}_t$. Ling and Li (1997) proposed an alternative portmanteau statistic for multivariate conditional heteroscedasticity. They defined the sample lag- h (transformed) residual autocorrelation as:

$$G(h) = \frac{\sum_{t=h+1}^T \left(\hat{\varepsilon}_t' \hat{H}_t^{-1} \hat{\varepsilon}_t - n \right) \left(\hat{\varepsilon}_{t-h}' \hat{H}_{t-h}^{-1} \hat{\varepsilon}_{t-h} - n \right)}{\sum_{t=h+1}^T \left(\hat{\varepsilon}_t' \hat{H}_t^{-1} \hat{\varepsilon}_t - n \right)^2}$$

Their test statistic is given by $LL(k) = T \sum_{h=1}^k G^2(h)$ and is asymptotically distributed as $\chi^2(k)$ under the null hypothesis. In the derivation of the asymptotic results, normality of the innovation process is not assumed.

¹⁶For more details see Francq and Raïssi (2007).

The statistic is thus robust with regard to the distribution choice. Tse and Tsui (2002) showed that there is a loss of information in the transformation $\hat{\varepsilon}_t' \hat{H}_t^{-1} \hat{\varepsilon}_t$ of the residuals and the test may suffer from a power reduction.

4 Multivariate GARCH Prediction

Forecasting is one of the main objectives of multivariate time series analysis. Predictions from multivariate GARCH models can be generated in a similar fashion to predictions from univariate GARCH models¹⁷. Indeed, for the univariate GARCH models, such as *CCC* model and principal component model, the predictions are generated from the underlying univariate GARCH models and then converted to the scale of the original multivariate time series by using the appropriate transformation. This section focuses on predicting from diagonal *BEKK* and *DCC* model.

To illustrate the prediction of conditional covariance matrix for multivariate GARCH models, consider the conditional variance equation for the diagonal *BEKK*(1,1,1) model:

$$H_t = C'C + A'\varepsilon_{t-1}\varepsilon_{t-1}'A + B'H_{t-1}B \quad (16)$$

where C , A and B are (3×3) matrices, C is upper triangular and A and B are diagonal matrices. The model (16) is estimated over the time period $t = 1, \dots, T$. Given the information at time T , the one-step-ahead prediction ($k = 1$) of conditional covariance matrix at time $T + 1$ is given by:

$$\begin{aligned} E_T(H_{T+1}) &= C'C + A'E_T(\varepsilon_T\varepsilon_T')A + B'E_T(H_T)B \\ &= C'C + A'\varepsilon_T\varepsilon_T'A + B'H_TB; \end{aligned}$$

when $k = 2$, it can be shown that

$$\begin{aligned} E_T(H_{T+2}) &= C'C + A'E_T(\varepsilon_{T+1}\varepsilon_{T+1}')A + B'E_T(H_{T+1})B \\ &= C'C + A'E_T(H_{T+1})A + B'E_T(H_{T+1})B \end{aligned}$$

where $E_T(H_{T+1})$ is obtained in the previous step. This procedure can be iterated to obtain $E_T(H_{T+k})$ for $k > 2$.

Let us consider the *DCC*(1,1)-*GARCH*(1,1) model given by (14), which can be written as:

$$\begin{cases} H_t = D_t R_t D_t, \quad D_t = \text{diag}(\sqrt{h_{iit}}) \text{ for } i = 1, \dots, n \\ h_t = c + \alpha \varepsilon_{t-1}^2 + \beta h_{t-1} \\ R_t = Q_t^{*-1} Q_t Q_t^{*-1} \\ Q_t = (1 - \theta_1 - \theta_2) \bar{Q} + \theta_1 u_{t-1} u_{t-1}' + \theta_2 Q_{t-1} \end{cases}$$

¹⁷For more details see Moon et al. (2008) and Hlouskova et al. (2009).

where $Q_t^{*-1} = \text{diag} \left(q_{ii}^{-1/2} \right)$, Q_t is the $(n \times n)$ symmetric positive definite matrix, θ_1 and θ_2 are non-negative scalar parameters satisfying $\theta_1 + \theta_2 < 1$ and \bar{Q} is the $(n \times n)$ unconditional covariance matrix composed from the standardized residuals resulting from the first step estimation. However, the k -step ahead forecast of a standard *GARCH(1,1)* and the *DCC* evolution process are given by:

$$\begin{cases} h_{t+k} = \sum_{i=0}^{k-2} c(\alpha + \beta)^i + (\alpha + \beta)^{k-1} h_{t+1} \\ \quad = \sum_{i=0}^{k-2} c(\alpha + \beta)^i + (\alpha + \beta)^k h_t \\ Q_{t+k} = (1 - \theta_1 - \theta_2) \bar{Q} + \theta_1 u_{t+k-1} u'_{t+k-1} + \theta_2 Q_{t+k-1} \end{cases} \quad (17)$$

where $E_t(u_{t+k-1} u'_{t+k-1}) = E_t(R_{t+k-1})$ and $R_{t+k} = Q_{t+k}^{*-1} Q_{t+k} Q_{t+k}^{*-1}$. Thus, the k -step ahead forecast of the correlation cannot be directly solved forward to provide a convenient method for forecasting. In examining methods to overcome this difficulty, two forecasts seem to be the most natural, each requiring a different set of approximations. The first technique proposed would be to generate the k -step ahead forecast of Q_t by making the approximation that $E_t(u_{t+1} u'_{t+1}) \approx Q_{t+1}$. Using this approximation, we then have the k -step ahead forecast of Q_t is

$$E_t(Q_{t+k}) = \sum_{i=0}^{k-2} (1 - \theta_1 - \theta_2) \bar{Q} (\theta_1 + \theta_2)^i + (\theta_1 + \theta_2)^{k-1} Q_{t+1} \quad (18)$$

for $i \in (1, \dots, k)$ and $R_{t+k} = Q_{t+k}^{*-1} Q_{t+k} Q_{t+k}^{*-1}$. An alternative approximation would be that $\bar{Q} \approx \bar{R}$ and that $E_t(Q_{t+1}) \approx E_t(R_{t+1})$. Using this approximation, we can forecast R_{t+k} directly using the relationship

$$E_t(R_{t+k}) = \sum_{i=0}^{k-2} (1 - \theta_1 - \theta_2) \bar{R} (\theta_1 + \theta_2)^i + (\theta_1 + \theta_2)^{k-1} R_{t+1} \quad (19)$$

In order to test which of these approximations performs better, Engle and shepard (2001) have conducted a Monte Carlo experiment. They have concluded that the forecast produced by solving forward for Q_{t+k} was more biased than the method for solving R_{t+k} forward which had better bias properties for almost all correlations and horizons. Also of interest is that both forecasts appear to be unbiased when the unconditional correlation is zero, and that they make essentially the same forecast when the unconditional correlation is zero.

While none of these two techniques significantly outperformed the other, it would seem that a logical choice for forecasting would be the method that directly solves forward R_{t+k} which appear easier to implement. Therefore, we will choose this second technique for further work.

5 Empirical application

5.1 The data

The data employed in this study are taken from Datastream and are the 4530 daily observations on the Nasdaq 100 (NAS), New York Stock Exchange composite (NYA) and Russell 3000 (RUA) stock returns over the period January 4, 1988 to December 21, 2005. The returns series denoted by r_t are calculated as $r_t = 100 \times \log(y_t/y_{t-1})$, where y_t is the price index. Models used are *Full-BEKK(1,1)*, *Diagonal BEKK(1,1)*, *CCC(1,1)-GARCH* and *DCC(1,1)-GARCH* where each of the univariate *GARCH* models estimated is a *GARCH(1,1)* and we focus our attention to the covariance matrix modelling.

Figures 1 and 2 give the plots of the daily price indices and daily stock returns. We can see that the market volatility is changing over time which suggests a suitable model for the data should have a time varying volatility structure as suggested by the GARCH model.

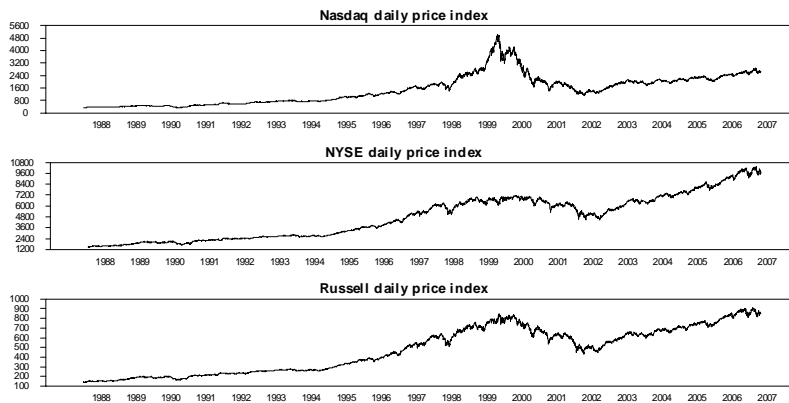


Figure 1. Daily price indices

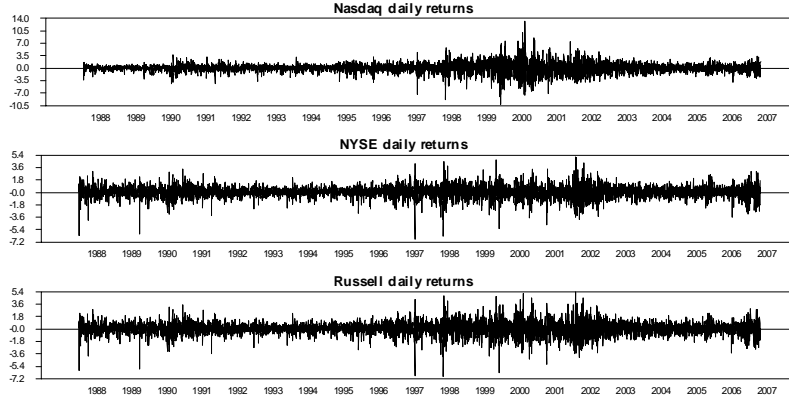


Figure 2. Daily stock returns

5.2 Estimation results

For the above mentioned indices, the sample mean of returns, standard deviation of returns, skewness and kurtosis coefficients as well as the Jarque-Bera and the Ljung-Box (univariate and multivariate version) tests are all reported in Table 1. Skewness is used to describe asymmetry from the normal distribution in a set of statistical data, taking two forms: positive or negative, it depends on whether data points are skewed to the left (negative coefficient) or to the right (positive coefficient) of the data average. Negative skewness means that there is a substantial probability of a big negative return, whereas positive one means that there is a greater than normal probability of a big positive return which indicates that the tail on the right side is longer than the left one and the bulk of the values lie to the left of the mean. The kurtosis measures the peakedness and fatness of the tails of a probability distribution. A fat tailed distribution has higher than normal chances of a big positive or negative realization. For all series, the returns distributions display positive skewness. Moreover, the data indicates high degree of excess kurtosis (leptokurtic), since the kurtosis coefficients are significantly larger than those of a normal distribution which is three. The returns series appear extremely non normal based on the Jarque-Bera test. The Ljung-Box test applied to the series and squared series, provides clear evidence against the hypothesis of serial independence of observations and indicates the existence of ARCH effect. The unit root tests of Phillips and Perron (1988), Kwiatkowski et al. (1992) and Dickey-Fuller Augmented (1979) reject the

stationarity hypothesis at both 5% and 1% significance levels for all series¹⁸.

The estimation of fractional integration parameter d based on the GPH procedure as well as the *CCC* and *DCC-GARCH* are given in Tables 2 and 3 respectively. The results reveal a clear evidence of long-range dependence for all stock returns, since d estimates are significantly positive implying covariance-stationarity of the process. These results are statistically significant and contrast with most of the studies of long memory in asset returns which have generally found weak or no evidence for long memory. However, Henry (2002) investigated long range dependence in nine international stock index returns and found evidence of long memory in four of them, the German, Japanese, South Korean and Taiwanese markets, but not for the markets of the UK, USA, Hong Kong, Singapore and Australia¹⁹. Furthermore, Serletis and Rosenberg (2009) analyzed daily data on four US stock market indices and concluded that US stock market returns display anti-persistence. This implies that the behaviour of stock returns is inconsistent with the efficient market hypothesis, which asserts that returns of a stock market are unpredictable from previous price changes (Narayan (2008) and Narayan and Smyth (2007)).

Moreover, we observe that the correlations across the three stock returns give strong evidence of time-varying correlations between them. The last column of Table 3 present the results of the estimation of the *DCC* parameters. we note that the estimate of θ_2 is statistically significant at 1% significance level meaning that the correlation is significantly time varying. The *DCC* parameters estimates imply a highly persistent correlation with $\hat{\theta}_1$ and $\hat{\theta}_2$, however, they satisfy the $0 < \hat{\theta}_1 + \hat{\theta}_2 < 1$ condition of stationarity. Thus, the model is mean reverting and the conditional correlation matrix is positive semi-definite. Apart the tables, we compute the Lagrange Multiplier statistic proposed by Tse (2000) for constant conditional correlation test for the trivariate model which is significant at the 5% level: its *p-value* is 0.0001. Thus, there is evidence against time-invariant correlations among the selected stock returns. In Figure 3²⁰, we observe a slow decay of the autocorrelation functions which indicate the presence of long memory behaviour. The plots of conditional variance of *BEKK-GARCH* models and *CCC-GARCH* are shown respectively in Figures 4 and 5.

¹⁸The results are not reported here to conserve space but are available from the authors upon request.

¹⁹See also Aydogan and Booth (1988).

²⁰R11, R12 and R13 is respectively the returns series of NAS, NYA and RUA.

Table 1. Basic properties of the distribution of returns series

	<i>NAS</i>	<i>NYA</i>	<i>RUA</i>
μ	0.586	0.502	0.521
σ	0.545	0.508	0.516
sk	2.065	2.882	2.705
K	9.651	21.082	18.431
JB	$5.305e + 003$ (0.000)	$3.117e + 004$ (0.000)	$2.314e + 004$ (0.000)
$Q(12)$	313.787 (0.000)	166.785 (0.000)	141.769 (0.000)
$Q^2(12)$	220.788 (0.000)	52.357 (0.000)	57.845 (0.000)
$Q_m(12)$	931.965 (0.000)	-	-
$Q_m^2(12)$	526.008 (0.000)	-	-

Notes: μ denotes the sample mean, σ is standard deviation, k is the Kurtosis (exc), sk is the Skewness coefficient, JB is the Jarque-Bera normality test, $Q(12)$ and $Q^2(12)$ are respectively the 12-th order Ljung-Box tests for serial correlation in the residuals and squared residuals. $Q_m(12)$ and $Q_m^2(12)$ are multivariate Ljung-Box version. The number in parenthese are the p-values.

Table 2. Estimation of the long memory parameter

	\hat{d}_{NAS}	\hat{d}_{NYA}	\hat{d}_{RUA}
<i>GPH</i>	0.301 (0.049)	0.206 (0.050)	0.245 (0.052)

Notes: GPH are the Geweke and Porter-Hudak (1983) procedure. The number in parenthese are standard errors.

Table 3. Estimation of the MGARCH model

Full BEKK(1,1,1)		Diagonal BEKK(1,1,1)		CCC-GARCH(1,1)		DCC-GARCH(1,1)	
c_{11}	0.095 (0.008)	c_{11}	0.084 (0.023)	c_{11}	0.007 (0.003)	c_{11}	0.005 (0.023)
c_{12}	0.089 (0.014)	c_{12}	0.068 (0.019)	c_{22}	0.012 (0.002)	c_{22}	0.008 (0.001)
c_{13}	0.099 (0.015)	c_{13}	0.072 (0.019)	c_{33}	0.012 (0.003)	c_{33}	0.007 (0.002)
c_{22}	0.062 (0.010)	c_{22}	0.046 (0.014)	α_{11}	0.066 (0.009)	α_{11}	0.064 (0.008)
c_{23}	0.064 (0.011)	c_{23}	0.041 (0.012)	α_{22}	0.075 (0.010)	α_{22}	0.040 (0.004)
c_{33}	0.0220 (0.001)	c_{33}	-0.007 (0.002)	α_{33}	0.078 (0.009)	α_{33}	0.064 (0.008)
α_{11}	0.103 (0.022)	α_{11}	0.214 (0.030)	β_{11}	0.932 (0.009)	β_{11}	0.934 (0.006)
α_{12}	0.040 (0.096)	α_{22}	0.198 (0.032)	β_{22}	0.901 (0.009)	β_{22}	0.948 (0.007)
α_{13}	-0.071 (0.118)	α_{33}	0.191 (0.028)	β_{33}	0.904 (0.013)	β_{33}	0.916 (0.010)
α_{21}	0.048 (0.023)	β_{11}	0.969 (0.009)	ρ_{12}	0.666 (0.011)	θ_1	0.042 (0.005)
α_{22}	0.151 (0.105)	β_{22}	0.972 (0.010)	ρ_{13}	0.666 (0.019)	θ_2	0.942 (0.008)
α_{23}	-0.113 (0.137)	β_{33}	0.973 (0.008)	ρ_{23}	0.429 (0.016)		
α_{31}	0.037 (0.023)						
α_{32}	-0.032 (0.123)						
α_{33}	0.131 (0.125)						
β_{11}	0.897 (0.013)						
β_{12}	-0.019 (0.069)						
β_{13}	0.007 (0.081)						
β_{21}	0.001 (0.013)						

Table 3. Continued

	Full BEKK(1,1,1)	Diagonal BEKK(1,1,1)	CCC-GARCH(1,1)	DCC-GARCH(1,1)
β_{22}	0.849 (0.072)			
β_{23}	-0.021 (0.082)			
β_{31}	0.011 (0.010)	-	-	-
β_{32}	0.072 (0.068)			
β_{33}	0.870 (0.083)			
AIC	-756.715	-743.369	-758.793	-760.490
BIC	-621.363	-511.044	-699.117	-698.454
HQ	-707.114	-768.169	-781.993	-787.756
$Q_m(12)$	646.262 [0.000]	487.183 [0.000]	141.769 [0.481]	151.019 [0.399]
$Q_m^2(12)$	423.541 [0.000]	361.060 [0.000]	85.242 [0.948]	99.283 [0.797]

Notes: The number in parenthese are standard error, ρ_{ij} is the correlation coefficient, AIC , BIC and HQ are respectively the Akaike, Bayesian and Hannan-Quinn information criterion. $Q_m(12)$ and $Q_m^2(12)$ are respectively the 12-th order multivariate Portmanteau tests for serial correlation in the standardized and squared standardized residuals (null Hypothesis: no serial correlation). The numbers in brackets are the p-values.

To check the goodness of fit of our model, we consider several diagnostic tests on the standardized residuals, the Ljung-Box test for the 12-th order serial autocorrelation and heteroskedasticity, when it is applied on the standardized residuals. We use the Jarque-Bera test, the skewness and the kurtosis coefficients to test the normality of standardized residuals. From Table 3, the multivariate portmanteau test reveal that the hypothesis of no residual autocorrelation is rejected in the residuals only for both *Full* and *Diagonal-BEKK(1,1,1)* models at 5% level of significance. From Table 4, we can see that, for most series, the hypothesis of uncorrelated standardized and squared standardized residuals is well supported, indicating that there is no statistically significant evidence of misspecification. The skewness and kurtosis coefficients indicate that the standardized residuals are still not normally

distributed, which is confirmed by the Jarque-Bera test. Finally, the hypothesis of no conditional heteroscedasticity are not rejected in the residuals for all the series at 5% level of significance.

Table 4. Diagnostic tests

	Full BEKK(1,1,1)			Diagonal BEKK(1,1,1)		
	<i>NAS</i>	<i>NYA</i>	<i>RUA</i>	<i>NAS</i>	<i>NYA</i>	<i>RUA</i>
$Q(12)$	166.771 (0.000)	270.155 (0.000)	6.996 (0.857)	43.283 (0.00202)	9.720 (0.640)	12.224 (0.427)
$Q^2(12)$	5.301 (0.947)	1.393 (0.999)	11.843 (0.458)	20.662 (0.055)	16.073 (0.187)	37.287 (0.021)
LM	6.204 (0.905)	2.776 (0.996)	11.760 (0.465)	7.236 (0.654)	2.939 (0.965)	3.197 (0.805)
sk	1.889	2.926	2.722	0.753	1.006	0.410
k	6.969	19.891	16.882	4.456	21.923	17.2967
JB	5449.493	3724.512	2725.128	380.193	31340	17700

	CCC-GARCH(1,1)			DCC-GARCH(1,1)		
	<i>NAS</i>	<i>NYA</i>	<i>RUA</i>	<i>NAS</i>	<i>NYA</i>	<i>RUA</i>
$Q(12)$	45.318 (0.0009)	0.838 (0.999)	1.066 (0.999)	8.735 (0.725)	5.711 (0.901)	5.390 (0.949)
$Q^2(12)$	8.758 (0.723)	0.012 (1.000)	0.013 (1.000)	4.491 (0.972)	1.593 (0.997)	1.690 (0.995)
LM	0.9395 (0.998)	0.0027 (1.000)	0.0035 (1.000)	0.229 (1.000)	1.176 (0.999)	1.197 (0.998)
sk	1.877	2.896	2.717	1.904	3.141	2.936
k	5.994	18.415	15.870	6.817	24.626	21.209
JB	4333.862	32283.720	24376.48	5283.913	55953.071	41956.369

Notes: k is the Kurtosis (exc), sk is the Skewness coefficient, JB is the Jarque-Bera normality test, $Q(12)$ and $Q^2(12)$ are respectively the 12-th order Ljung-Box tests for serial correlation in the standardized and squared standardized residuals (Null Hypothesis: no autocorrelation). LM is the test for ARCH Effects (Null Hypothesis: no ARCH effects).

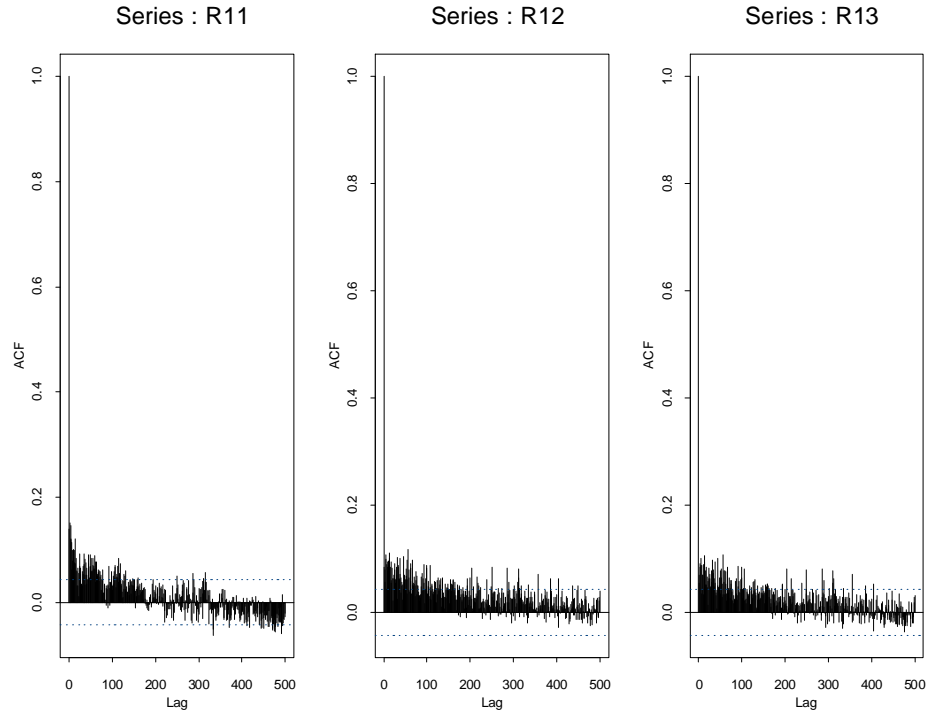


Figure 3. Autocorrelation of returns series

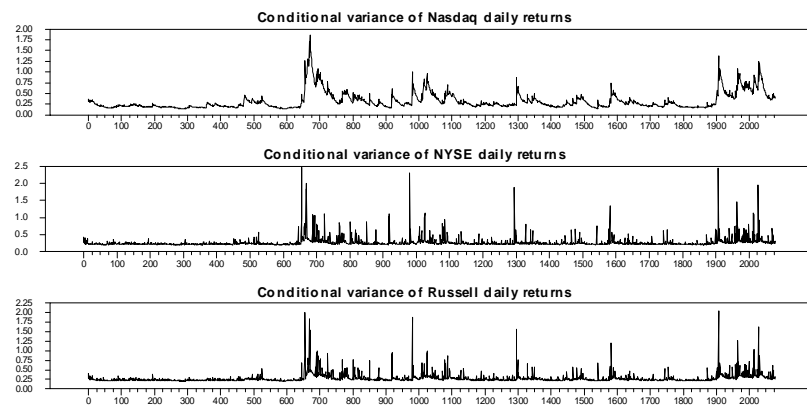


Figure 4. Conditional variance of returns series in BEKK(1,1,1)

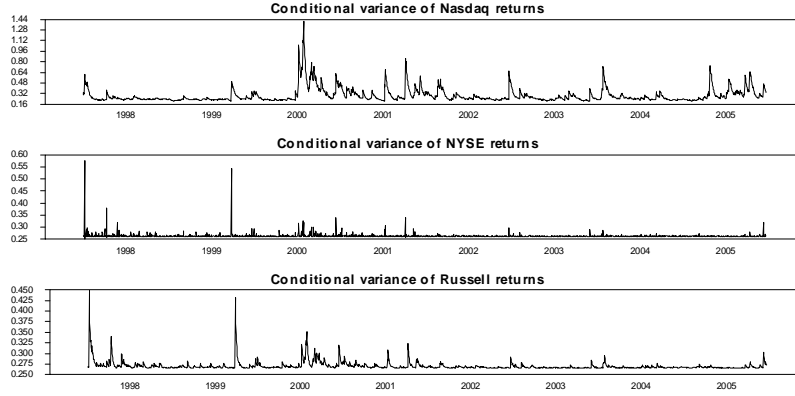


Figure 5. Conditional variance of returns series in CCC models

5.3 Forecasting performance of estimated models

The *DCC* model seems better with respect to the other models in terms of information criteria. So, in order to assess the out-of-sample forecasting performance for diagonal *BEKK*, *CCC* and *DCC* models, we still use root mean square error (RMSE) and mean absolute error (MAE) as two criteria for comparison. we have selected an out-of-sample forecast data set using the last 1000 observations of the original data. We have re-estimated the models adding a new observation and obtaining the k -day-step ahead forecasts for $k = 1, 3$ and 5. The results are shown in Table 5. They appear to show a trend that the forecasting errors are proportionate to the forecasted periods. Moreover, we draw clear inference to the effect that they all appeared to be more accurate in *DCC* than in the other models, regardless of what criterion is adopted. This seems to be consistent not only in RMSE but also in MAE (the *DCC* model has the lowest RMSE and MAE). In addition, predicting results of *CCC* perform even worse than diagonal *BEKK*. The results successfully provide evidence in favour of the predictive superiority of the *DCC* model against the diagonal *BEKK* and *CCC*. We can thus conclude that the forecasts of the *DCC* modelling are significantly better than those of the other model.

Table 5. Out-of-sample forecast

		Diag BEKK	CCC	DCC
<i>RMSE</i>	1	0.506	0.783	0.429
	3	0.545	0.785	0.445
	5	0.611	0.790	0.478
<i>MAE</i>	1	0.359	0.452	0.318
	3	0.368	0.457	0.329
	5	0.389	0.460	0.361

Conclusion

The aim of the present paper is to study the dynamic modelling of the US stock returns. We considered multivariate GARCH framework to model the time-varying covariance matrices of a process exhibiting a long term dependence and be used to produce out-of-sample forecasts. In particular, we examined the persistence phenomenon in the first conditional moment of daily stock returns; the robustness of the results was also investigated by considering that its innovations was generated by a multivariate GARCH process. As illustration, we applied our models to the trivariate systems. The estimated parameters show that the returns series are characterized by long memory behaviour and time-varying correlations. These results are statistically significant and contrast with most of the studies of long memory in the returns series showing weak or no evidence for long term dependence. Using daily returns of Nasdaq 100, New York Stock Exchange composite and Russell 3000, the results successfully provide evidence that *DCC* model outperforms the other ones in estimating and forecasting covariance matrices for out-of-sample analysis.

Acknowledgement

I would like to thank Anne Peguin and Mohamed Boutahar for their useful comments on an earlier version of this paper. I'm especially grateful to the anonymous reviewer and the editor (P.K. Narayan) for their detailed comments and numerous suggestions that helped to improve the paper. Any errors and or omissions are, however, are my own.

References

- [1] Aydogan, K. and G. G. Booth (1988), "Are there long cycles in common stock returns?", *Southern Economic Journal* 55, 141-149.
- [2] Barkoulas, John T., Baum, C. F. and N. Travlos (2000), "Long memory in the Greek stock market", *Applied Financial Economics* 10, 2, 177-184.
- [3] Bauwens, L., Laurent, S. and J. Rombouts (2006), "Multivariate GARCH models: A survey", *Journal of Applied Econometrics* 21, 79-109.
- [4] Bera A.K. and S. Kim (2002), "Testing constancy of correlation and other specifications of the BGARCH model with an application to international equity returns", *Journal of Empirical Finance* 9, 171-195.
- [5] Beran, J. (1994), "Statistics for long memory processes", vol 61 of *Mono-graphs on Statistics and Applied Probability*, Chapman & Hall/CRC, Boca Raton.
- [6] Billio, M. and M. Caporin (2006), "A generalized dynamic conditional correlation model for portfolio risk evaluation", *Working Paper No. 53/WP/2006*, Department of Economics, Ca Foscari University of Venice.
- [7] Bollerslev, T. (1990), "Modeling the coherence in short-run nominal exchange rates: a multivariate generalized ARCH model", *Review of Economics and Statistics* 72, 498-505.
- [8] Bollerslev, T., Engle, R. and J.M. Wooldridge, (1988), "A capital asset pricing model with time varying covariances", *Journal of Political Economy* 96, 116-131.
- [9] Brooks C., Burke S.P. and G. Persaud (2003), "Multivariate GARCH models: software choice and estimation issues", *Journal of Applied Econometrics* 18, 725-734.
- [10] Cajueiro, D. O. and B. M. Tabak (2004), "Evidence of long range dependence in Asian equity markets: The role of liquidity and market restrictions", *Physica A* 342, 656-664.
- [11] Cajueiro, D. O. and B. M. Tabak (2005), "Possible causes of long-range dependence in the Brazilian stock market", *Physica A: Statistical Mechanics and its Applications*, 345, 635-645.

- [12] Cappiello, L., Engle, R. F. and K. Sheppard (2006), "Asymmetric dynamics in the correlations of global equity and bond returns", *Journal of Financial Econometrics* 4, 537-572.
- [13] Cavalcante, J. and A. Assaf (2004), "Long-range dependence in the returns and volatility of the Brazilian stock market", *European Review of Economics and Finance* 3, 5-22.
- [14] Ding, Z., Granger, C.W.J. and R.F. Engle (1993), "A long memory property of stock market returns and a new model", *Journal of Empirical Finance* 1, 83-106.
- [15] Doukhan, P., Oppenheim, G., and M.S. Taqqu (2003), "Theory and applications of long-range dependence", *Birkhäuser, Basel*.
- [16] Duchesne, P. and S. Lalancette (2003), "On Testing for Multivariate ARCH Effects in Vector Time Series Models", *The Canadian Journal of Statistics* 31, 3, 275-292.
- [17] Engle, R.F. (2002), "Dynamic conditional correlation: A simple class of multivariate GARCH models", *Journal of Business and Economic Statistics* 20, 339-350.
- [18] Engle, R.F. and K. Sheppard (2001), "Theoretical and empirical properties of dynamic conditional correlation multivariate GARCH", *Mimeo, UCSD*.
- [19] Engle, R. F. and K. F. Kroner (1995), "Multivariate simultaneous GARCH", *Econometric Theory*, 11, 122-150.
- [20] Francq, C. and H. Raïssi (2007), "Multivariate Portmanteau Test For Autoregressive Models with Uncorrelated but Nonindependent Errors", *Journal of Time Series Analysis* 28, 3, 454-470.
- [21] Geweke, J., and S. Porter-Hudak (1983), "The estimation and application of long-memory time series models", *Journal of Time Series Analysis* 4, 221-238.
- [22] Gil-Alana, L.A. (2007), "A multivariate long memory model for the specification of real output in the US, the UK, and Canada", *International Journal of Business and Economics* 2, 135-146.
- [23] Gil-Alana, L.A. (2003), "A fractional multivariate long memory model for the US and the Canadian real output", *Economics Letters* 81, 355-359.
- [24] Goetzmann, W.N. (1993), "Patterns in three centuries of stock market prices", *The Journal of Business* 66, 2, 249-270.

- [25] Granger, C. W. J. and R. Joyeux (1980), "An introduction to long memory time series models and fractional differencing", *Journal of Time Series Analysis* 1, 15-29.
- [26] Greene, M. T. and B. D. Fielitz (1977), "Long-term dependence in common stock returns", *Journal of Financial Economics* 4, 3, 339-349.
- [27] Henry, Ólan T. (2002), "Long memory in stock returns: Some international evidence", *Applied Financial Economics* 12, 10, 725-729.
- [28] Hlouskova, J., K. Schmidheiny and M. Wagner (2009), "Multistep predictions for multivariate GARCH models: Closed form solution and the value for portfolio management", *Journal of Empirical Finance* 16, 330-336.
- [29] Hosking, J. R. M. (1981b), "Lagrange-tests of multivariate time series models", *Journal of the Royal Statistical Society B* 43, 219-30.
- [30] Hosking, J. R. M. (1980), "The multivariate portmanteau statistic", *Journal of American Statistical Association* 75, 602-608.
- [31] Hosoya, Y. (1996), "The quasi-likelihood approach to statistical inference on multiple time-series with long-range dependence", *Journal of Econometrics* 73, 217-236.
- [32] Huang, B. N. and C. W. Yang (1999), "An examination of long-term memory using the intraday stock returns", *Technical Report 99-03*, Clarion University of Pennsylvania.
- [33] Hurst, H. E. (1951), "The long-term storage capacity of reservoirs", *Trans. Am. Soc. Civil Eng.*, 118.
- [34] Jegadeesh, N. (1990), "Evidence of predictable behavior of security returns", *Journal of Finance* 45, 881-898.
- [35] Kim, M., Nelson, C. and Startz, R. (1991), "Mean reversion in stock prices? a reappraisal of the empirical evidence", *Review of Economic Studies* 58, 515-528.
- [36] Limam, I. (2003), "Is long memory a property of thin stock markets? International evidence using Arab countries", *Review of Middle East Economics and Finance* 1, 3, 251-266.
- [37] Ling, S. and W. K. Li (1997), "Diagnostic checking of nonlinear multivariate time series with multivariate ARCH errors", *Journal of Time Series Analysis* 18, 5, 447-464.
- [38] Lo, A.W. (1991), "Long-term memory in stock market prices", *Econometrica*, 59, 5, 1279-1313.

- [39] Lobato, I.N. (1999), "A semiparametric two-step estimator in a multivariate long memory model", *Journal of Econometrics*, 90, 129-153.
- [40] Longin F. and B. Solnik (1995), "Is the Correlation in International Equity Returns Constant: 1960-1990?", *Journal of International Money and Finance*, 14, 3-26.
- [41] Luceño, A. (1996), "A fast likelihood approximation for vector general linear processes with long series: Application to fractional differencing", *Biometrika* 83, 603-614 .
- [42] Lux, T. (1996), "Long-term stochastic dependence in financial prices: Evidence from the German stock market". *Applied Economics Letters* 3, 701-706.
- [43] Mills, T. C. (1993), "Is there long-term memory in UK stock returns?", *Applied Financial Economics*, 3, 4, 303-306.
- [44] Moon, G. H., Yu, W. C. and C. H. Hong (2008), "Hedging performance of commodity futures: out-of-sample evaluation of multivariate GARCH models", *Kyonggi University*.
- [45] Narayan, P.K. (2008), "Do shocks to G7 stock prices have a permanent effect? Evidence from panel unit root tests with structural change", *Mathematics and Computers in Simulation* 77, 369-373.
- [46] Narayan, P.K. and R. Smyth (2007), "Mean reversion versus random walk in G7 stock prices evidence from multiple trend break unit root tests", *International Financial Markets, Institutions and Money* 17, 152-166.
- [47] Nelson, D.B. and C.Q. Cao (1992), "Inequality constraints in the univariate GARCH model", *Journal of Business and Economic Statistics* 10, 229-235.
- [48] Nielsen, M. (2004), "Efficient inference in multivariate fractionally integrated time series models", *Econometrics Journal* 7, 63-97.
- [49] Pelletier, D. (2006), "Regime switching for dynamic correlations", *Journal of Econometrics* 131, 445-473.
- [50] Peters, E. (1991), "Chaos and order in the capital markets. A new view of cycles, prices, and market volatility", *John Wiley and Sons*.
- [51] Robinson, P. M. (1995b), "Gaussian semi-parametric estimation of long range dependence", *The Annals of Statistics*, 23, 1630-1661
- [52] Sadique, S. and P. Silvapulle (2001), "Long-term memory in stock market returns: international evidence", *International Journal of Finance & Economics* 6, 1, 59-67.

- [53] Serletis, A. and A. A. Rosenberg (2009), "Mean reversion in the US stock market", *Chaos, Solitons & Fractals* 40, 2007–2015.
- [54] Sheppard, K. (2003), "Multi-step estimation of Multivariate GARCH models", *University of California at San Diego*.
- [55] Silvennoinen, A. and T. Teräsvirta (2009), "Modelling multivariate autoregressive heteroskedasticity with the double smooth transition conditional correlation GARCH model", *Journal of Financial Econometrics* 7, 373–411.
- [56] Silvennoinen, A. and T. Teräsvirta (2008), "Multivariate GARCH models", To appear in T. G. Andersen, R. A. Davis, J.-P. Kreiss and T. Mikosch, eds. *Handbook of Financial Time Series*. New York: Springer. *SSE/EFI Working Paper Series in Economics and Finance N° 669*.
- [57] Silvennoinen, A. (2008), "Numerical aspects of the estimation of multivariate GARCH models", *QFRC Research Paper*, University of Technology, Sydney.
- [58] Sowell, F. (1989b), "Maximum likelihood estimation of fractionally integrated time series models", Unpublished manuscript.
- [59] Tolvi, J. (2003), "Long Memory in a small stock market", *Economics Bulletin* 7, 3, 1–13.
- [60] Tolvi, J. (2003), "Long memory and outliers in stock market returns", *Applied Financial Economics* 13, 7, 495–502.
- [61] Tsay, W.J. (2007), "Maximum Likelihood Estimation of Stationary Multivariate ARFIMA Process", *IEAS Working Paper N°. 07-A011*.
- [62] Tse, Y. K. (2000), "A test for constant correlations in a multivariate GARCH model", *Journal of Econometrics* 98, 107–127.
- [63] Tse, Y.K. and A. K.C. Tsui (2002), "A multivariate generalized autoregressive conditional heteroscedasticity model with time-varying correlations", *Journal of Business and Economic Statistics* 20, 351–362.
- [64] Volos, C.C. and F. M. Siokis (2006), "Long range dependence in stock market returns", *Applied Financial Economics* 16, 1331–1338.
- [65] Zhuang, Y., Green, C. J. and P. Maggioni (2000), "The great rebound, the great crash, and persistence in British stock prices", *Technical Report 00/11*, Loughborough University, Loughborough.